



# Aplikasi Analisis Komponen Utama dan Analisis Gerombol pada Varietas Tanaman Hias Krisan (*Chrysanthemum morifolium* R.) di Kota Tomohon

Ranni A. Pangkey<sup>1</sup>, Yohanes A. R. Langi<sup>1</sup>, Hanny A. H. Komalig<sup>1\*</sup>

<sup>1</sup>Jurusan Matematika–Fakultas Matematika dan Ilmu Pengetahuan Alam–Universitas Sam Ratulangi Manado, Indonesia

\*Corresponding author : [hanov07@yahoo.com](mailto:hanov07@yahoo.com)

## ABSTRAK

Kota Tomohon dikenal sebagai produsen bunga (kembang) di Provinsi Sulawesi Utara. Bunga lokal Kota Tomohon adalah bunga krisan yang memiliki banyak varietas. Penelitian ini bertujuan untuk mengeksplorasi, menentukan komponen utama dan mengelompokkan data varietas krisan di Kota Tomohon. Sampel yang di ambil adalah 10 varietas krisan di Kota Tomohon dengan 14 variabel indikator. Penelitian ini dilakukan pada bulan Maret sampai Juni 2018. Metode yang digunakan adalah Analisis Komponen Utama dan Analisis Gerombol Non Hirarki untuk mereduksi data dan mengelompokkan data berdasarkan kemiripannya. Hasil eksplorasi data dengan Analisis Komponen Utama didapat 3 komponen utama yaitu  $KU_1$ ,  $KU_2$ , dan  $KU_3$  yang mewakili 78,8% variabilitasnya dan terbentuk 3 kelompok dengan menggunakan Analisis Gerombol yaitu kelompok pertama memuat varietas krisan Elora, Kulo dan Ririh, kelompok kedua memuat varietas krisan Solinda Pelangi, Salzieta, Kineta dan Arosuka Pelangi, kelompok ketiga memuat varietas krisan Pasopati, Merahayani dan Limeron.

## INFO ARTIKEL

Diterima : 17 Juli 2018

Diterima setelah revisi : 23 Juli 2018

Tersedia online : 30 Juli 2018

## Kata Kunci:

Varietas Krisan

Analisis Komponen Utama

Analisis Gerombol

## 1. PENDAHULUAN

Kota Tomohon adalah salah satu kota di Provinsi Sulawesi Utara, Indonesia. Luas Kota Tomohon berdasarkan keputusan UU RI Nomor 10 Tahun 2003 sekitar 11.420 Ha [1]. Kota Tomohon terletak di ketinggian kira-kira 900-1100 meter dari permukaan laut (mdpl). Suhu di Kota Tomohon pada waktu siang mampu mencapai 30° celsius dan 18°-22° celsius pada malam hari [2]. Kota Tomohon kini dikenal sebagai produsen bunga (kembang) di Provinsi Sulawesi Utara.

Krisan (*Chrysanthemum*) adalah jenis tumbuhan yang berbunga yang sering ditanam sebagai tanaman hias pekarangan atau bunga petik yang memiliki bermacam-macam jenis. Beberapa varietas krisan tumbuh di Kota Tomohon diantaranya 2 varietas krisan lokal yaitu Kulo dan Ririh.

Banyaknya varietas dari tanaman hias krisan ini dibutuhkan pengelompokan objek agar dapat mengetahui objek mana yang berada dalam kelompok yang sama. Analisis gerombol (*Cluster Analysis*) merupakan suatu teknik analisis multivariat yang mempunyai tujuan utama untuk melakukan pengelompokan objek-objek pengamatan menjadi beberapa kelompok berdasarkan karakteristik yang dimilikinya [3].

Dalam pengelompokan objek ini perlu diketahui komponen utamanya untuk mewakili objek yang diteliti. Komponen utama tersebut mampu mempertahankan sebagian besar informasi yang diukur menggunakan sedikit variabel yang menjadi komponen utamanya saja [4].

Analisis komponen utama dan analisis gerombol merupakan analisis tertua dalam analisis peubah ganda sehingga beberapa peneliti telah menggunakan metode dari analisis ini.

Penelitian sebelumnya dengan menggunakan Analisis Komponen Utama pada kasus produksi pertanian dan perkebunan di wilayah Bolaang Mongondow tahun 2008 dengan menggabungkan dua gugus data peubah ganda [5] dan penelitian menggunakan Analisis Gerombol pada pengelompokan Sekolah Dasar di Siau, dimana penggerombolan dilakukan dengan menentukan komponen utama terlebih dahulu sebagai dasar penentuan gerombol dan melanjutkan ke analisis gerombol [6].

Berdasarkan latar belakang tersebut, peneliti tertarik untuk melakukan studi lebih lanjut tentang bagaimana menentukan komponen utama dan mengelompokkan data dari varietas tanaman hias krisan yang ada di Kota Tomohon.

## 2. ANALISIS KOMPONEN UTAMA

Analisis komponen utama digunakan untuk identifikasi variabel baru yang mendasari data peubah ganda, mereduksi jumlah himpunan peubah banyak dan saling berkorelasi dengan mempertahankan sebanyak mungkin keragaman data tersebut, dan menghilangkan variabel-variabel asal yang tidak memberi informasi yang penting. Menurut [7], proses pembentukan Komponen Utama sebagai berikut :

- Misalkan matriks  $X$  merupakan data pengamatan yang berukuran  $n \times p$ , dengan  $n =$  objek dan  $p =$  variabel

$$X_{n \times p} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \vdots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

Kemudian dicari  $X'X$  untuk menjadi matriks baru dalam analisis data selanjutnya.

- Dari matriks  $X'X$ , dicari matriks korelasi ( $r_{pp}$ )

$$R_{p \times p} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1p} \\ r_{21} & r_{22} & \dots & r_{2p} \\ \dots & \dots & \vdots & \dots \\ r_{p1} & r_{p2} & \dots & r_{pp} \end{bmatrix}$$

- Menentukan nilai *eigen* dari matriks korelasi, misalkan  $\lambda_1, \lambda_2, \dots, \lambda_p$  dengan  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$
- Menentukan vektor *eigen* yang bersesuaian dengan setiap nilai *eigen*, untuk  $j = 1, 2, \dots, p$ , yaitu  $v_j = v_{j1}, v_{j2}, \dots, v_{jp}$
- Berdasarkan matriks *eigen*, maka komponen utama yang terbentuk adalah:

$$KU_1 = v_{j1}'X_j = v_{11}X_1 + v_{21}X_2 + \dots + v_{j1}X_p$$

$$KU_2 = v_{j2}'X_j = v_{12}X_1 + v_{22}X_2 + \dots + v_{j2}X_p \quad (1)$$

$$\vdots$$

$$KU_p = v_{jp}'X_j = v_{1p}X_1 + v_{2p}X_2 + \dots + v_{jp}X_p$$

- Kriteria yang digunakan untuk mentukan beberapa komponen yang dibentuk adalah kriteria persen varian 75%.

## 3. ANALISIS GEROMBOL

Metode analisis gerombol membutuhkan suatu ukuran ketakmiripan (jarak) yang didefinisikan untuk setiap pasang objek yang akan dikelompokkan. Pengukuran jarak yang digunakan dalam penelitian ini yaitu jarak *Euclidean* dapat dirumuskan sebagai berikut:

$$d_{i,j} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (2)$$

Dengan:  $d_{i,j}$  = jarak antara objek  $i$  dengan objek  $k$

$x_{ik}$  = nilai objek  $i$  pada variabel ke- $k$

$x_{jk}$  = nilai objek  $j$  pada variabel ke- $k$

$p$  = banyaknya variabel yang diamati

Prosedur *cluster* atau pengelompokan data dapat dilakukan dengan metode *K-means*. Menurut [8] proses pengelompokan dengan *k-means* adalah :

- Tentukan jumlah *cluster*
- Alokasikan data ke dalam *cluster* secara random
- Hitung *centroid* (rata-rata) dari data yang ada di masing-masing *cluster*
- Alokasikan masing-masing data ke *centroid* (rata-rata) terdekat

- Kembali ke langkah 3 apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*, jika pusat *cluster* sudah tidak berubah maka proses selesai.

## 4. METODOLOGI PENELITIAN

Penelitian dilaksanakan pada bulan Maret sampai Mei 2018 di Kota Tomohon. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Dinas Pertanian, Perikanan dan Perkebunan Kota Tomohon dan Balai Penelitian Tanaman Hias (BALITHI). Data yang diambil adalah deskripsi morfologi 10 varietas tanaman hias krisan yang tumbuh di Kota Tomohon dengan 14 variabel indikator yang memiliki skala yang berbeda. Adapun 10 varietas krisan tersebut yaitu Solinda Pelangi, Salzieta, Pasopati, Merahayani, Limeron, Kineta, Elora, Arosuka Pelangi, Kulo dan Ririh.

Tahapan analisis yang dilakukan adalah sebagai berikut:

- Analisis Komponen Utama dimana dilakukan untuk menentukan jumlah dan bentuk persamaan komponen utama. Hasil analisis komponen utama dijadikan dasar untuk menentukan jumlah kelompok yang akan di analisis gerombol.
- Analisis Gerombol yaitu diambil skor dari setiap komponen utama dijadikan data analisis untuk penggerombolan karena data asal memiliki satuan skala yang berbeda maka perlu di standarisasi dengan mengambil skor komponen-komponen utama yang terbentuk. Penggerombolan dengan metode *K-means* dengan menentukan jumlah kelompok yang akan dibentuk.

## 5. HASIL DAN PEMBAHASAN

### 5.1. Analisis Komponen Utama

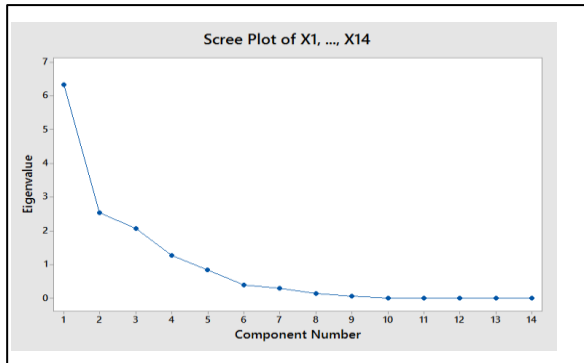
Setelah data dianalisis korelasi selanjutnya membentuk komponen utama untuk menentukan jumlah kelompok yang akan digunakan dalam analisis gerombol. Pertama menentukan nilai eigen dan vektor eigen dari matriks korelasi. Nilai eigen dari setiap komponen utamanya dituangkan dalam tabel 1.

Tabel 1. Proporsi Nilai Eigen

KU	Nilai Eigen	Proporsi	Kumulatif
1	6.4278	0.459	0.459
2	2.5452	0.182	0.641
3	2.0583	0.147	0.788
4	1.2745	0.091	0.879
5	0.8313	0.059	0.938
6	0.3654	0.026	0.964
7	0.2867	0.02	0.985
8	0.1363	0.01	0.995
9	0.0745	0.005	1
10	0	0	1
11	0	0	1
12	0	0	1
13	0	0	1
14	0	0	1

Dari tabel 1 diambil beberapa komponen utama yang bisa mewakili data variabel-variabel asal. Jika hanya diambil satu komponen utama yang pertama proporsi nilai eigennya hanya 45,9% sehingga dianggap belum dapat menggambarkan keragaman total, tetapi bila diakumulasikan proporsi dari 3 komponen utama yaitu  $KU_1, KU_2, KU_3$  menghasilkan

proporsi 0,788 yang artinya ketiga komponen utama tersebut bisa mewakili 78,8% dari total variabilitasnya. Nilai eigen pada tabel 1 dituangkan dalam *scree plot* pada gambar 1.



Gambar 1. *Scree Plot* data Komponen Utama

Gambar 1 menunjukkan nilai eigen dari komponen utama pertama ( $KU_1$ ) sampai komponen utama terakhir ( $KU_{14}$ ) yang semakin mengecil yaitu dari kiri atas ke kanan bawah.

Tabel 2. Matriks Komponen Utama

Variabel	$KU_1$	$KU_2$	$KU_3$
$X_1$	0.0797168	0.1625054	-0.293592
$X_2$	-0.243062	-0.001046	-0.4673625
$X_3$	0.2691888	-0.05543	-0.4854265
$X_4$	-0.0063302	0.3512217	0.44094682
$X_5$	0.0196448	0.5505222	-0.2764611
$X_6$	-0.0359141	0.5534888	-0.2365893
$X_7$	0.1884479	0.2927081	0.18092726
$X_8$	0.3442388	-0.206349	-0.176743
$X_9$	0.3652646	-0.129161	-0.0808773
$X_{10}$	0.3857732	-0.032098	-0.0763758
$X_{11}$	0.2609615	0.1785088	-0.0311619
$X_{12}$	0.3763129	0.0905072	0.13731643
$X_{13}$	0.3284486	0.1090967	0.01128179
$X_{14}$	-0.3264118	-0.199518	-0.1913021

Tabel 3. Skor Antar 3 Komponen Utama

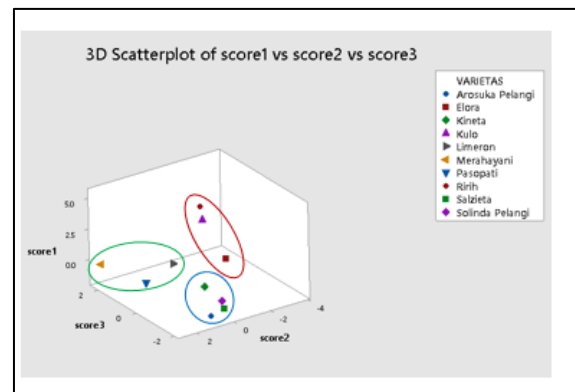
Skor 1	Skor 2	Skor 3
-1.029634	-0.07856	-1.554343561
-1.586343	-0.05872	-1.642396044
-0.678051	1.839989	0.566714413
-0.323974	2.650169	2.175702539
-1.244643	-0.69119	1.636864734
-1.423122	-0.57968	-0.141439953
-1.672923	-3.22271	1.317695074
-1.459284	0.963502	-1.958999554
4.054177	-0.51535	-0.159930699
5.363797	-0.30744	-0.239866695

Tabel 2 merupakan vektor eigen dari 3 komponen utama dengan masing-masing skor dituangkan dalam

tabel 3. Dari matriks komponen utama pada tabel 2, jika diturunkan dalam bentuk persamaan matematis adalah sebagai berikut :

$$\begin{aligned}
 KU_1 &= 0,080X_1 - 0,243X_2 + 0,269X_3 - 0,006X_4 + \\
 &\quad 0,020X_5 - 0,036X_6 + 0,188X_7 + 0,344X_8 + \\
 &\quad 0,365X_9 + 0,386X_{10} - 0,261X_{11} + 0,376X_{12} \\
 &\quad + 0,328X_{13} - 0,326X_{14} \\
 KU_2 &= 0,163X_1 - 0,001X_2 - 0,055X_3 + 0,351X_4 + \\
 &\quad 0,551X_5 + 0,553X_6 + 0,293X_7 - 0,206X_8 - \\
 &\quad 0,129X_9 - 0,032X_{10} + 0,179X_{11} + 0,091X_{12} \\
 &\quad + 0,109X_{13} - 0,200X_{14} \\
 KU_3 &= -0,294X_1 - 0,467X_2 - 0,485X_3 + 0,441X_4 - \\
 &\quad 0,276X_5 - 0,237X_6 + 0,181X_7 - 0,177X_8 - \\
 &\quad 0,081X_9 - 0,076X_{10} - 0,031X_{11} + 0,137X_{12} \\
 &\quad + 0,011X_{13} - 0,191X_{14}
 \end{aligned}$$

Plot skor dari tabel 3 jika dilihat dari 3 komponen utama pertama yaitu  $KU_1$ ,  $KU_2$ , dan  $KU_3$  disajikan dalam gambar 2



Gambar 2. Plot Skor 3 Komponen Utama

Dari gambar 2 memperlihatkan 3 komponen utama yaitu  $KU_1$ ,  $KU_2$ , dan  $KU_3$  yang membentuk 3 kelompok. Jumlah kelompok ini yang nantinya akan menjadi acuan pengelompokan varietas krisan dalam analisis gerombol.

## 5.2. Pengelompokan Varietas Krisan Dengan Menggunakan Analisis Gerombol

Selanjutnya dilakukan analisis gerombol dengan menggunakan skor komponen-komponen utama yang ada pada tabel 3 sebagai variabel indikatornya. Analisis gerombol varietas krisan menggunakan algoritma *K-means* dengan jumlah gerombol 3 dapat dilihat pada tabel 4, tabel 5, tabel 6 dan tabel 7.

Tabel 4. Hasil Akhir Analisis Gerombol

	Number of Observations	Within Cluster Sum of Squares	Average Distance from Centroid	Maximum Distance from Centroid
Gerombol 1	3	7.184	1.448	2.166
Gerombol 2	4	1.472	0.528	0.917
Gerombol 3	3	3.106	0.994	1.249

Tabel 4 merupakan hasil akhir dari analisis gerombol dengan *software minitab 17* memperlihatkan jumlah observasi dari 3 gerombol yang terbentuk dimana gerombol pertama memuat 3 varietas krisan, gerombol kedua memuat 4 varietas

krisan, gerombol ketiga memuat 3 varietas krisan. Jumlah kuadrat dari gerombol (*within cluster sum of squares*) dan jarak rata-rata dari pusat gerombol (*average distance from centroid*) bernilai rendah maka gerombol tersebut memiliki tingkat kemiripan yang besar, begitupun sebaliknya. Ini berarti gerombol 2 lebih mirip dari gerombol 3 dan gerombol 1, dan gerombol 3 lebih mirip dari gerombol 1. Ditampilkan dalam tabel 6 nilai jarak maksimum dari centroid yang lebih tinggi yaitu pada gerombol 1 menunjukkan observasi di gerombol 1 terletak lebih jauh dari pusat gerombol.

**Tabel 5. Nilai Rata-rata dari Setiap Gerombol**

Var	Gerombol 1	Gerombol 2	Gerombol 3
Skor 1	1.0183	-0.5422	-0.2954
Skor 2	-0.8453	0.0386	0.7938
Skor 3	0.2133	-0.9231	1.0175

Pada tabel 5, terlihat bahwa, pada gerombol 1 memiliki nilai yang lebih tinggi dari gerombol 2, tetapi gerombol 3 memiliki nilai yang paling tinggi diantara kedua gerombol tersebut.

**Tabel 6. Jarak Rata-rata Antar Gerombol**

	Gerombol 1	Gerombol 2	Gerombol 3
Gerombol 1	0	2.1231	2.2492
Gerombol 2	2.1231	0	2.0969
Gerombol 3	2.2492	2.0969	0

Pada tabel 6, jarak terbesar yaitu pada gerombol 1 dengan gerombol 3 yang menunjukkan 2 gerombol tersebut memiliki perbedaan yang besar.

**Tabel 7. Penempatan Varietas ke dalam Gerombol**

Varietas	Gerombol
Solinda Pelangi	2
Salzietta	2
Pasopati	3
Merahayani	3
Limeron	3
Kineta	2
Elora	1
Arosuka Pelangi	2
Kulo	1
Ririh	1

Tabel 7 memperlihatkan penempatan gerombol varietas yaitu gerombol pertama terdiri dari 3 varietas krisan yaitu Elora, Kulo dan Ririh, gerombol kedua terdiri dari 4 varietas krisan yaitu Solinda Pelangi, Salzietta, Kineta dan Arosuka Pelangi, dan gerombol ketiga terdiri dari 3 varietas krisan yaitu Pasopati, Merahayani dan Limeron.

## 6. KESIMPULAN

- 1) Eksplorasi data dengan komponen utama dalam menentukan jumlah gerombol dari 10 varietas krisan dengan 14 variabelnya adalah 3 komponen utama pertama yaitu  $KU_1, KU_2, KU_3$  yang mewakili 78,8% variabilitasnya. Dari plot skor 3 komponen utama terbentuk 3 gerombol sebagai dasar penentuan nilai  $k$  untuk proses penggerombolan menggunakan metode non hirarki  $k$ -means.
- 2) Analisis gerombol  $k$ -means dengan nilai  $k = 3$  menghasilkan 3 gerombol tanaman hias krisan yang memiliki tingkat kemiripan yaitu gerombol pertama terdiri dari 3 varietas krisan yaitu Elora, Kulo dan Ririh, gerombol kedua terdiri dari 4 varietas krisan yaitu Solinda Pelangi, Salzietta, Kineta dan Arosuka Pelangi, dan gerombol ketiga terdiri dari 3 varietas krisan yaitu Pasopati, Merahayani dan Limeron.
- 3) Varietas krisan lokal Kota Tomohon yaitu Kulo dan Ririh berada pada satu gerombol yang sama yaitu pada gerombol 1 bersama dengan varietas krisan Elora.

## REFERENSI

- [1] Undang-Undang RI, 2003. Pembentukan Kabupaten Minahasa Selatan dan Kota Tomohon di Provinsi Sulawesi Utara. Jakarta, Jakarta.
- [2] Badan Pusat Statistika Kota Tomohon. 2017. Tomohon Dalam Angka 2017. BPS Tomohon, Tomohon.
- [3] Johnson, R.A., and D. W. Wichern. 2007. *Applied Multivariate Statistical Analysis*. 6<sup>th</sup> Ed. Pearson Prentice Hall, New Jersey.
- [4] Johnson R.A and D.W. Wichern. 2002. *Applied Multivariate Analysis*. 5<sup>th</sup> Ed. Prentice Hall Inc, New Jersey.
- [5] Sunarsi, A. H. Komalig, N. Nainggolan. 2014. Penggunaan Analisis Komponen Utama Dalam Penggabungan Data Peubah Ganda pada Kasus Produksi Pertanian dan Perkebunan di Wilayah Bolaang Mongondow Tahun 2008. *Jurnal De Cartesian* 3(2):1 – 8.
- [6] Matahari, M. J. S. Kekenusa, Y. A. R. Langi. 2015. Pengelompokan Sekolah Dasar Di Siau berdasarkan Indikator Mutu Sekolah dengan Menggunakan Analisis Gerombol. *Jurnal De Cartesian* 4(2):188 - 195
- [7] Rencher, A.C. 2001. *Methods of Multivariate Analysis*. Second Edition. A Wiley-Interscience Publication, New York.
- [8] MacQueen, J. B., 1967. Some methods for Classification and Analysis of Multivariate Observation, Proceedings of 5<sup>th</sup> Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley.

**Ranni A. Pangkey** ([pangkeyranni@gmail.com](mailto:pangkeyranni@gmail.com))



Lahir dan tinggal di Taratara, Kecamatan Tomohon Barat, Kota Tomohon. Menempuh pendidikan tinggi Program Studi Matematika, FMIPA, Universitas Sam Ratulangi Manado. Tahun 2018 adalah tahun terakhir ia menempuh studi. Makalah ini merupakan hasil penelitian skripsinya yang dipublikasikan.

**Hanny A. H. Komalig** ([hanoy07@yahoo.com](mailto:hanoy07@yahoo.com))



Pada tahun 1990, memperoleh gelar Insinyur di Fakultas Pertanian, Universitas Sam Ratulangi Manado. Kemudian ia mengikuti *Basic Sciences Bridging Program* bidang Matematika di Institut Teknologi Bandung pada tahun 1992 selama 2 tahun. Selanjutnya memperoleh gelar Magister Sains bidang Statistika di Institut Pertanian Bogor tahun 1998.

Dan pada tahun 2008 memperoleh gelar Doktor di Universitas Airlangga Surabaya untuk bidang Matematika Modeling. Menjadi dosen di departemen Matematika, FMIPA, Universitas Sam Ratulangi pada tahun 2000 sampai sekarang. Bidang keahlian yang ditekuni yaitu diantaranya; Statistika Multivariat Nonlinier dan Reduksi Dimensi.

**Yohanes A. R. Langi** ([varlangi@unsrat.ac.id](mailto:varlangi@unsrat.ac.id))



Pada tahun 1994, memperoleh gelar sarjana di Departemen Matematika, Fakultas MIPA Universitas Kristen Tomohon. Gelar magister bidang Biometrika Hutan diperoleh dari Departemen Biometrika, Institut Pertanian Bogor, pada tahun 2007. Menjadi dosen di departemen Matematika, FMIPA, Universitas Sam Ratulangi

Manado sejak tahun 2005 sampai sekarang dengan bidang keahlian yang ditekuni diantaranya; Proses Stokastik, Perancangan Percobaan, dan Biometrika.